substitution of claims 15-39 are believed to put the present application in better form for review by the Examiner.

Respectfully submitted,

Mitchell B. Wasson
Reg. 27,408

March 8, 2002

HOFFMAN, WASSON & GITLER, PC
2361 Jefferson Davis Highway
Suite 522
Arlington, VA 22202
(703) 415-0100

Attorney's Docket: A-7200.PAM/cat

The distribution function of speech is modeled as a GMM of time-correlated samples, leading to a distribution function for the speech spectral amplitude $s(f)$ as shown in Equation 2, where $\delta(s)$ is a one-sided Dirac delta function. The first term on the right hand side (RHS) of Equation 2 represents a signal of zero power, thus capturing the possibility that no signal of interest is present. The components of the summation in the second term on the RHS of Equation 2 are the components of the GMM model for the speech distribution function.

Equation 2

$$f_s(s) = (1 - q_S)\delta(s) + q_S\{2s\sum_i \frac{a_i}{\rho_i}\ Exp(-s^2/\rho_i)\}$$

This speech model has two sets of frequency band dependent parameters which are dynamically updated during the processing, $\{P_s(f)\}$ and $\{q_s(f)\}$. The first is the *a priori* PSD of the speech, assuming that a speech signal is present at the frequency and time of interest. The second ~~parameter~~ is the *a priori* probability of a speech signal being present at that frequency and time. The speech distribution function also has a number of added parameters, $\{a_I\} = \{a_1, a_2, \ldots a_N\}$ and $\{\rho_i^\circ\} = (\rho_1^\circ, \rho_2^\circ, \ldots \rho_N^\circ\}$. The $\{a_i\}$ are the weights of the N Gaussian components of the GMM, and the $\{\rho_i^\circ\}$ are the powers of each component when the speech PSD is normalized to $P_s(f) = 1$. In practice, $P_s(f)$ and $\{\rho_i^\circ\}$ are combined into a parameter set denoted as $\{\rho_i(f)\}$, where $\rho_i(f) = \rho_i^\circ\ P_s(f)$.

While both the $P_s(f)$ and $q_s(f)$ are dynamically updated during the processing, the $\{a_i\}$ and ~~are~~ $\{\rho_i^\circ\}$ are determined from prior "training" to optimize processing results as averaged over a representative body of training data. This may typically be done by minimizing the mean-squared-error (MSE) between noise free signals and the results from processing noisy input signals based on those signals by mixing with varying types and levels of interfering noise. The present invention may typically use five GMM components (denoted GMM5). However, more or less than five components can be employed. In addition, the $\{a_i\}$ may be further parameterized by the values of other key quantities, including but not limited to signal-to-noise ratio (SNR), which are adaptively and dynamically updated throughout the processing. This may typically be done by determining different GMM model parameter values (the $\{a_i\}$ and $\{\rho_i^\circ\}$) versus SNR based on training for different input SNRs, and interpolating between these model parameter values based on the adaptively estimated input SNR during the processing. One prior training of a GMM5 leads to a model for the speech distribution as shown in Figure 1 for q_s = 0.5. Also shown is the corresponding distribution function for a Gaussian speech model with q_s = 1. For presentation purposes, the vertical axis is actually the distribution function for speech spectral power, which is simply $f(s^2/P_s)$, and the horizontal axis is (s^2/P_s)[illegible].

Noise PSD updating is mainly based on the following. Given *a priori* distribution functions for the noise and speech spectral amplitudes, and a new measurement of the noisy signal spectral amplitude, $r(f)$, a determination is made as to a best

The form of this estimator is depicted in Figures 4a and 4b. In these figures, the vertical axis is $(\langle s^2 | r \rangle / P_N)^{1/2}$, and the horizontal axis is $(r^2/P_N)^{1/2}$. GMM5 results are given for different SNRs, a nominal speech distribution function at q_s = 0.5, and as compared with a Gaussian speech model at q_s = 1.0, and also an extended Gaussian modes at q_s = 0.5. GMM5 results are in solid lines and Gaussian models are shown as dashed lines.

In a manner similar to the previous explanation, the speech spectral amplitude can also be estimated as follows.

Equation 13 <<<<<~~the equation~~ has ~~been changed~~>>>>>

$$\langle \bar{s} | \bar{r} \rangle = \frac{q_S \sum_i \frac{a_i S_i}{(1+S_i)^2} Exp[(\frac{r^2}{P_N})(\frac{S_i}{1+S_i})]}{(1-q_S)+q_S \sum_i \frac{a_i}{1+S_i} Exp[(\frac{r^2}{P_N})(\frac{S_i}{1+S_i})]} \bar{r}$$

Change the vector in these 3 places

Note that in the special case with only one GMM component in the speech distribution function, and also with q_s = 1, the above expression reduces to a conventional Wiener filter.

For a typical set of GMM parameters, and at q_s = *0.5*, and for different SNRs, the form of this estimator is shown in Figures 5a and 5b, where it is also compared with a Wiener filter at q_s = *1.0*, and also with an extended Wiener filter based on a Gaussian speech model but with q_s = *0.5*. In the figures, the vertical axis is $\langle s | r \rangle / (P_N)^{1/2}$, and the horizontal axis is $(r^2/P_N)^{1/2}$.

It is further noted that the availability of separate estimates for both the speech spectral amplitude $\langle s | r \rangle$ and the